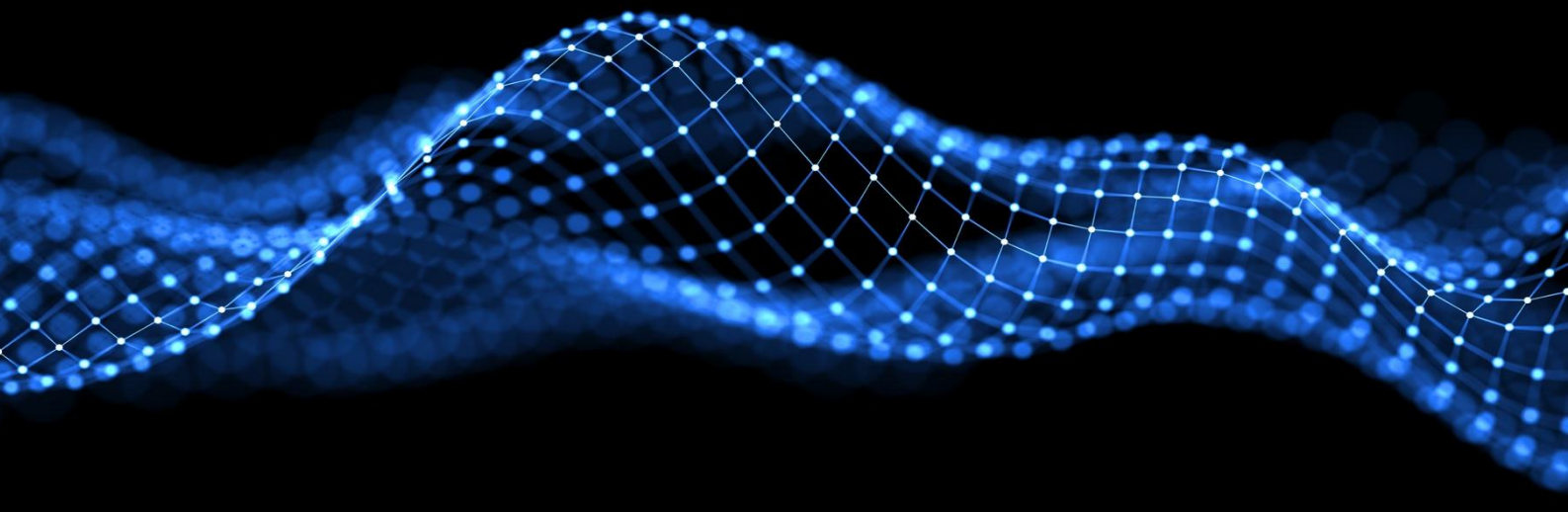


Pure Storage ObjectEngine



PRODUCT BRIEF

The Architect's View



Pure Storage ObjectEngine is a scale-out S3-compatible gateway that optimises data written using the S3 API. Data is de-duplicated and divided into blocks that more effectively serves the requirements of data protection solutions using object storage as a backup target. Pure Storage combines ObjectEngine and FlashBlade together to create a solution that competes with data de-duplication appliances.

Background

Object and file storage solutions are increasingly being used as targets for storing backup data. Public cloud-based solutions such as AWS S3 and Microsoft Azure Blob Storage provide relatively cheap, uniform pricing for data based on capacity and access requirements.

Pure Storage identified that customers were using FlashBlade (their scale-out object and file solution) as a target for secondary backup data. This wasn't an original use-case for the platform, however customers had determined that "fast restore" could be achieved when using FlashBlade as a backup repository.

Public cloud repositories are delivered as a service and as such, obfuscate much of the technology that underlies the solution. In general, this doesn't cause a problem for users of the platform, however when using cloud object storage for backup data, there are some challenges to overcome.

Public cloud storage providers charge on volume of data and don't pass on any hardware optimisation savings that can be achieved. This means with types of data like that generated by backup, there is a high rate of duplication that customers have to pay for.

Similarly, FlashBlade has no inherent de-duplication capability. The technology was originally developed for AI and ML workloads where the overhead of including de-duplication could have affected performance. As a result,

FlashBlade has the same challenges dealing with highly de-duplicatable data as public cloud.

Public cloud offerings store data based on individual immutable objects and generally don't offer the ability to easily modify content. If backup systems simply wrote data as one amorphous blob of content, then significant amounts of data would be retained on the platform, even after some backup content had expired.

To best use object storage, either on-premises or in public cloud, data protection solutions need a way to optimise for the requirements of backup data and the design restrictions of the object storage platform.

What is ObjectEngine?

ObjectEngine from Pure Storage is a hardware and software appliance solution that provides the interface between backup software and object storage platforms. The technology has been developed from the acquisition by Pure Storage of StorReduce in August 2018.

StorReduce developed a gateway solution that runs on public cloud. As input, the software appliance exposes an S3 interface. At the back-end, StorReduce writes de-duplicated and optimised data to a public storage account like AWS S3. The gateway functionality implements global de-duplication and sharding of data. The result is a saving in storage costs from de-duplication while implementing a level of granularity that ensures data isn't retained on public cloud longer than necessary.

Pure Storage has taken the StorReduce software, initially sold as a virtual appliance in AWS, and packaged the solution as an on-premises de-duplication engine. ObjectEngine runs on physical appliances and can scale-out to match the throughput and performance demands of enterprise data protection.

ObjectEngine can write data to both FlashBlade on-premises or AWS S3 public cloud, providing a fast restore and long-term archive solution for secondary data.

Customer Business Value

ObjectEngine provides the capability to realise significant storage savings on secondary backup data. Pure Storage quotes figures of 8:1 reduction, however customers may see significantly better results, depending on their data and backup regime in use.

Looking at how the data protection industry has developed, de-duplication engines introduced in the 2000's (most notably Data Domain owned by EMC) changed the way in which backup data was stored on-premises. They allowed HDD systems to replace tape and provide high levels of operational efficiency and reliability.

ObjectEngine enables customers to replace their storage targets within existing backup infrastructure. The solution itself isn't a complete backup platform but designed to provide efficient data storage in the way the de-duplication devices of the 2000s did.

Reference Information

Further details on ObjectEngine and FlashBlade can be found in the following Architecting IT blog posts:

- [Pure //Accelerate 2016 – FlashBlade](#) (Published 29 March 2016)
- [Pure Storage Acquires StorReduce](#) (Published 22 August 2018)
- [Pure Storage – Seeding the NVMe Market](#) (Published 5 September 2018)
- [Pure Storage ObjectEngine for Flash-based Backup](#) (Published 19 March 2019)

Further details on ObjectEngine and FlashBlade can be found with the following Storage Unpacked podcasts:

- [#109 – An Overview of ObjectEngine with Brian Schwarz](#) (Published 19 July 2019)
- [Soundbytes #008: FlashBlade 2.0 With Rob Lee at Pure Accelerate](#) (Published 14 June 2017)

Information in this document is provided for guidance only and is offered for use without warranty or other guarantees. Brookend Ltd accepts no liability whatsoever for decisions made from the basis of this information. All statements made are expressed in good faith, based on briefings and other publicly available information. Where available, links are provided to sources used for the production of this document. Pure Storage has been or is an existing client of Brookend Ltd.

Copyright © 2019 Brookend Ltd, all rights reserved. The exploding box logo is a trademark of Brookend Ltd, all other logos and trademarks acknowledged to their respective owners. DocID: BRKBD0108. Last revised 22 July 2019. Revision 1.0.

Market Positioning

ObjectEngine provides Pure Storage two immediately obvious benefits. Firstly, FlashBlade can be marketed with a wider set of use cases, including both ML/AI and secondary data on the same platform. Second, Pure Storage can appeal to a wider group of enterprise customers that want to take their storage solutions from only a single vendor. Pure can now offer to replace both primary and secondary storage solutions within the data centre.

Caveats

ObjectEngine is a gateway and as such, stores the metadata that translates backup "objects" from the data protection software, into de-duplicated storage on the object storage platform. The object data is not self-describing and needs the ObjectEngine platform (either in software or hardware) to access the data. Customers need to validate how Pure Storage protects the ObjectEngine metadata, in case of hardware failure or site loss. Without this, the data stored becomes unusable.

As ObjectEngine is acting only as a storage target, customers will need to replicate or move existing backups into the platform using data movers built into the backup solution itself. This could be done through attrition over time, or by proactive migration.